LUD 5793.1

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
APPLICATION FOR LETTERS PATENT

TITLE: **ISOLATED NUCLEIC ACID MOLECULES ENCODING CANCER ASSOCIATED ANTIGENS, THE ANTIGENS PER SE, AND USES THEREOF**

EXPRESS MAIL

Mailing Label Number EV 331558628 US

Date of Deposit December 4, 2003

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" Service under 37 CFR 1.10 on the date indicated above and is addressed to the Mail Stop: Patent Applications, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313

Fani Malikouzakis
(Typed or printed name of person mailing paper or fee)

Fani Malikouzakis
(Signature of person mailing paper or fee)

Norman Hanson
Reg. No. 30,946
Fulbright & Jaworski L.L.P.
666 Fifth Avenue
New York, N.Y. 10103
(212) 318-3148
(212) 318-3400 (fax)

ISOLATED NUCLEIC ACID MOLECULES ENCODING CANCER ASSOCIATED ANTIGENS, THE ANTIGENS PER SE, AND USES THEREOF

RELATED APPLICATIONS

[0001] This application is a continuation-in-part of Serial No. 60/430,869, filed December 4, 2002, which is a continuation-in-part of Serial No. 10/181,663, filed November 29, 2000 which is a continuation-in-part of Serial No. 09/602,362, filed June 22, 2000 which is a continuation in part of Serial No. 09/451,739, filed November 30, 1999, all of which are incorporated by reference in their entirety.

FIELD OF THE INVENTION

[0002] This invention relates to antigens associated with cancer, the nucleic acid molecules encoding them, as well as the uses of these.

BACKGROUND AND PRIOR ART

[0003] It is fairly well established that many pathological conditions, such as infections, cancer, autoimmune disorders, etc., are characterized by the inappropriate expression of certain molecules. These molecules thus serve as "markers" for a particular pathological or abnormal condition. Apart from their use as diagnostic "targets", i.e., materials to be identified to diagnose these abnormal conditions, the molecules serve as reagents which can be used to generate diagnostic and/or therapeutic agents. A by no means limiting example of this is the use of cancer markers to produce antibodies specific to a particular marker. Yet another non-limiting example is the use of a peptide which complexes with an MHC molecule, to generate cytolytic T cells against abnormal cells.

[0004] Preparation of such materials, of course, presupposes a source of the reagents used to generate these. Purification from cells is one laborious, far from sure method of doing so. Another preferred method is the isolation of nucleic acid molecules which encode a particular marker, followed by the use of the isolated encoding molecule to express the desired molecule.

[0005] Two basic strategies have been employed for the detection of such antigens, in e.g., human tumors. These will be referred to as the genetic approach and the biochemical approach. The genetic approach is exemplified by, e.g., dePlaen, et al., *Proc. Natl. Sci. USA*,

85:2275 (1988), incorporated by reference. In this approach, several hundred pools of plasmids of a cDNA library obtained from a tumor are transfected into recipient cells, such as COS cells, or into antigen-negative variants of tumor cell lines which are tested for the expression of the specific antigen. The biochemical approach, exemplified by, e.g., O. Mandelboim, et al., *Nature*, 369:69 (1994) incorporated by reference, is based on acidic elution of peptides which have bound to MHC-class I molecules of tumor cells, followed by reversed-phase high performance liquid chromography (HPLC). Antigenic peptides are identified after they bind to empty MHC-class I molecules of mutant cell lines, defective in antigen processing, and induce specific reactions with cytotoxic T-lymphocytes. These reactions include induction of CTL proliferation, TNF release, and lysis of target cells, measurable in an MTT assay, or a 51Cr release assay.

[0006] These two approaches to the molecular definition of antigens have the following disadvantages: first, they are enormously cumbersome, time-consuming and expensive; and second, they depend on the establishment of cytotoxic T cell lines (CTLs) with predefined specificity.

[0007] The problems inherent to the two known approaches for the identification and molecular definition of antigens is best demonstrated by the fact that both methods have, so far, succeeded in defining only very few new antigens in human tumors. See, e.g., van der Bruggen, et al., *Science*, 254:1643-1647 (1991); Brichard, et al., *J. Exp. Med.*, 178:489-495 (1993); Coulie, et al., *J. Exp. Med.*, 180:35-42 (1994); Kawakami, et al., *Proc. Natl. Acad. Sci. USA*, 91:3515-3519 (1994).

[0008] Further, the methodologies described rely on the availability of established, permanent cell lines of the cancer type under consideration. It is very difficult to establish cell lines from certain cancer types, as is shown by, e.g., Oettgen, et al., *Immunol. Allerg. Clin. North. Am.*, 10:607-637 (1990). It is also known that some epithelial cell type cancers are poorly susceptible to CTLs in vitro, precluding routine analysis. These problems have stimulated the art to develop additional methodologies for identifying cancer associated antigens.

[0009] One key methodology is described by Sahin, et al., *Proc. Natl. Acad. Sci. USA*, 92:11810-11913 (1995), incorporated by reference. Also, see U.S. Patent No. 5,698,396. These references are incorporated by reference. To summarize, the method

involves the expression of cDNA libraries in a prokaryotic host. (The libraries are secured from a tumor sample). The expressed libraries are then immunoscreened with absorbed and diluted sera, in order to detect those antigens which elicit high titer humoral responses. This methodology is known as the SEREX method ("Serological identification of antigens by Recombinant Expression Cloning"). The methodology has been employed to confirm expression of previously identified tumor associated antigens, as well as to detect new ones. See the above referenced patent and Sahin, et al., supra, as well as Crew, et al., *EMBO J.*, 144:2333-2340 (1995), also incorporated by reference.

[0010] This methodology has been applied to a range of tumor types, including those described by Sahin, et al., supra, and Pfreundschuh, supra, as well as to esophageal cancer (Chen, et al., *Proc. Natl. Acad. Sci. USA*, 94:1914-1918 (1997)); lung cancer (Güre, et al., *Cancer Res.*, 58:1034-1041 (1998)); colon cancer (Serial No. 08/948, 705 filed October 10, 1997) incorporated by reference, and so forth. Among the antigens identified via SEREX are the SSX2 molecule (Sahin, et al., *Proc. Natl. Acad. Sci. USA*, 92:11810-11813 (1995); Tureci, et al., *Cancer Res.*, 56:4766-4772 (1996); NY-ESO-1 Chen, et al., *Proc. Natl. Acad. Sci. USA*, 94:1914-1918 (1997); and SCP1 (U.S. Patent No. 6,043,084) incorporated by reference. Analysis of SEREX identified antigens has shown overlap between SEREX defined and CTL defined antigens. MAGE-1, tyrosinase, and NY-ESO-1 have all been shown to be recognized by patient antibodies as well as CTLs, showing that humoral and cell mediated responses do act in concert.

[0011] It is clear from this summary that identification of relevant antigens via SEREX is a desirable aim. The inventors have applied this methodology and have identified several new antigens associated with cancer, as detailed in the description which follows.

BRIEF DESCRIPTION OF THE FIGURES

[0012] Figures 1-3, inclusive, show that NY-BR-1 is processed to peptides that are recognized by naturally occurring, CD8+ cells.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

EXAMPLE 1

[0013] The SEREX methodology, as described by, e.g. Sahin, et al., *Proc. Natl. Acad. Sci. USA*, 92:11810-11813 (1995); Chen, et al., *Proc. Natl. Acad. Sci. USA*, 94:1914-1918 (1997), and U.S. Patent No. 5,698,396, all of which are incorporated by reference. In brief, total RNA was extracted from a sample of a cutaneous metastasis of a breast cancer patient (referred to as "BR11" hereafter), using standard CsCl guanidine thiocyanate gradient methodologies. A cDNA library was then prepared, using commercially available kits designed for this purpose. Following the SEREX methodology referred to supra, this cDNA expression library was amplified, and screened with either autologous BR11 serum which had been diluted to 1:200, or with allogeneic, pooled serum, obtained from 7 different breast cancer patients, which had been diluted to 1:1000. To carry out the screen, serum samples were first diluted to 1:10, and then preabsorbed with lysates of E. coli that had been transfected with naked vector, and the serum samples were then diluted to the levels described supra. The final dilutions were incubated overnight at room temperature with nitrocellulose membranes containing phage plaques, at a density of 4-5000 plaque forming units ("pfus") per 130 mm plate.

[0001] Nitrocellulose filters were washed, and incubated with alkaline phosphatase conjugated, goat anti-human Fcγ secondary antibodies, and reactive phage plaques were visualized via incubation with 5-bromo-4-chloro-3-indolyl phosphate and nitroblue tetrazolium.

[0015] This procedure was also carried out on a normal testicular cDNA library, using a 1:200 serum dilution.

[0016] A total of 1.12x106 pfus were screened in the breast cancer cDNA library, and 38 positive clones were identified. With respect to the testicular library, 4x105 pfus were screened, and 28 positive clones were identified.

[0017] Additionally, 8x105 pfus from the BR11 cDNA library were screened using the pooled serum described. Of these, 23 were positive.

[0018] The positive clones were subcloned, purified, and excised to forms suitable for insertion in plasmids. Following amplification of the plasmids, DNA inserts were

evaluated via restriction mapping (EcoRI-XbaI), and clones which represented different cDNA inserts were sequenced using standard methodologies.

[0019] If sequences were identical to sequences found in GenBank, they were classified as known genes, while sequences which shared identity only with ESTs, or were identical to nothing in these data bases, were designated as unknown genes. Of the clones from the breast cancer library which were positive with autologous serum, 3 were unknown genes. Of the remaining 35, 15 were identical to either NY-ESO-1, or SSX2, two known members of the CT antigen family described supra, while the remaining clones corresponded to 14 known genes. Of the testicular library, 12 of the clones were SSX2.

[0020] The NY-ESO-1 antigen was not found, probably because the commercial library that was used had been size fractionated to have an average length of 1.5 kilobases, which is larger than full length NY-ESO-1 cDNA which is about 750 base pairs long.

[0021] With respect to the screening carried out with pooled, allogeneic sera, four of the clones were NY-ESO-1. No other CT antigens were identified. With the exception of NY-ESO-1, all of the genes identified were expressed universally in normal tissue.

[0022] A full listing of the isolated genes, and their frequency of occurrence follows, in tables 1, 2 and 3. Two genes were found in both the BR11 and testicular libraries, i.e., poly (ADP-ribose) polymerase, and tumor suppression gene ING1. The poly (ADP-ribose) polymerase gene has also been found in colon cancer libraries screened via SEREX, as is disclosed by Scanlan, et al., *Int. J. Cancer*, 76:652-58 (1998) when the genes identified in the screening of the BR11 cDNA library by autologous and allogeneic sera were compared, NY-ESO-1 and human keratin.

Table 1. SEREX-defined genes identified by autologous screening of BR11 cDNA library

Gene group	No. of clones	Comments	Expression
CT genes	10	NY-ESO-1	tumor, testis
	5	SSX2	tumor, testis
Non-CT genes	5	Nuclear Receptor Co-Repressor	ubiquitous
	4	Poly(ADP-ribose) polymerase	ubiquitous
	2	Adenylosuccinatelyase	ubiquitous
	2	cosmid 313 (human)	ESTs: muscle, brain, breast
	1	CD 151 (transmembrane protein)	ubiquitous

	1	Human HRY Gen	RT-PCR: multiple normal tissues
	1	Alanyl-t-RNA-Synthetase	ubiquitous
	1	NAD(+) ADP-Ribosyltransferase	ubiquitous
	1	Human keratin 10	ESTs: multiple normal tissues
	1	Human EGFR kinase substrate	ubiquitous
	1	*ING 1* Tumor suppressor gene	RT-PCR: multiple normal tissues
	1	Unknown gene, NCI_CGAP_Pr12 cDNA clone	ESTs: pancreas, liver, spleen, uterus
	1	Unknown gene	ESTs: multiple normal tissues
	1	Unknown gene	RT-PCR: multiple normal tissues

Table 2. SEREX-defined genes identified by allogeneic screening of BR11 cDNA library

Gene group	No. of clones	Comments	Expression
CT genes	4	NY-ESO-1	tumor, testis
Non-CT genes	6	zinc-finger helicase	ESTs: brain, fetal heart, total fetus
	4	Acetoacetyl-CoA-thiolase	ubiquitous
	3	KIAA0330 gene	ESTs: multiple normal tissues
	2	U1snRNP	ubiquitous
	1	Human aldolase A	ubiquitous
	1	Retinoblastoma binding protein 6	ESTs: tonsils, fetal brain, endothelial cells, brain
	1	α2-Macroglobulin receptor associated protein	ubiquitous
	1	Human Keratin 10	ESTs: multiple normal tissues

Table 3. SEREX-defined genes identified by screening of a testicular cDNA library with BR11 serum

Gene group	No. of clones	Comments	Expression
CT genes:	12	SSX2	tumor, testis
Non-CT genes:	3	Rho-associated coiled-coil forming protein	ubiquitous
	3	Poly(ADP-ribose) polymerase	ubiquitous
	3	Gene from HeLa cell, similar to TITIN	ubiquitous
	2	Gene from parathyroid tumor	RT-PCR: multiple normal tissues
	1	Transcription termination factor I-interacting peptide 21	ubiquitous

	1	Gene from fetal heart	ESTs: multiple normal tissues
	1	*ING 1* tumor suppressor gene	RT-PCR: multiple normal tissues
	1	KIAA0647 CdnA	ESTs: multiple normal tissues
	1	KIAA0667 cDNA	ESTs: multiple normal tissues

EXAMPLE 2

[0023] The mRNA expression pattern of the cDNAs identified in example 1, in both normal and malignant tissues, was studied. To do this, gene specific oligonucleotide primers were designed which would amplify cDNA segments 300-600 base pairs in length, using a primer melting temperature of 65-70° C. The primers used for amplifying MAGE-1,2,3 and 4, BAGE, NY-ESO-1, SCP1, and SSX1, 2, 3, 4 and 5 were known primers, or were based on published sequences. See Chen, et al. supra; Tureci, et al., *Proc. Natl. Acad. Sci. USA*, 95:5211-16 (1998). Gure, et al., *Int. J. Cancer*, 72:965-71 (1997); Chen, et al., *Proc. Natl. Acad. Sci. USA*, 91:1004-1008 (1994); Gaugler, et al., *J. Exp. Med.*, 179:921-930 (1994), dePlaen, et al., *Immunogenetics*, 40:360-369 (1994), all of which are incorporated by reference. RT-PCR was carried out for 35 amplification cycles, at an annealing temperature of 60° C. Using this RT-PCR assay, the breast cancer tumor specimen was positive for a broad range of CT antigens, including MAGE-1,3 AND 4, BAGE, SSX2, NY-ESO-1 and CT7. The known CT antigens SCP-1, SSX1, 4 and 5 were not found to be expressed.

[0024] An additional set of experiments were carried out, in which the seroreactivity of patient sera against tumor antigens was tested. Specially, ELISAs were carried out, in accordance with Stockert, et al., *J. Exp. Med.*, 187:1349-1354 (1998), incorporated by reference, to determine if antibodies were present in the patient sera. Assays were run for MAGE-1, MAGE-3, NY-ESO-1, and SSX2. The ELISAs were positive for NY-ESO-1 and SSX2, but not the two MAGE antigens.

EXAMPLE 3

[0025] Two clones (one from the breast cancer cDNA library and one from the testicular library), were identified as a gene referred to as ING1, which is a tumor suppressor gene candidate. See Garkavtsev, et al., *Nature*, 391:295-8 (1998), incorporated by reference. The sequence found in the breast cancer library, differed from the known sequence of ING1 at six residues, i.e., positions 818, 836, 855, 861, 866 and 874. The sequence with the six

variants is set forth at SEQ ID NO: 1. The sequence of wild type ING1 is set out at SEQ ID NO: 2.

[0026] To determine if any of these differences represented a mutation in tumors, a short, PCR fragment which contained the six positions referred to supra was amplified from a panel of allogeneic normal tissue, subcloned, amplified, and sequenced following standard methods.

[0027] The results indicated that the sequences in the allogeneic tissues were identical to what was found in tumors, ruling out the hypothesis that the sequence differences were a tumor associated mutation. This conclusion was confirmed, using the testicular library clone, and using restriction analysis of ING1 cDNA taken from normal tissues. One must conclude, therefore, that the sequence information provided by Garkavtsev, et al., supra, is correct.

EXAMPLE 4

[0028] Additional experiments were carried out to determine whether genetic variations might exist in the 5′ portion of the ING1 gene, which might differ from the 5′ portion of the clone discussed supra (SEQ ID NO: 1). In a first group of experiments, attempts were made to obtain full length ING1 cDNA from both the breast tumor library, and the testicular library. SEQ ID NO: 1 was used as a probe of the library, using standard methods.

[0001] Four clones were isolated from the testicular library and none were isolated from the breast cancer library. The four clones, following sequencing, were found to derive from three transcript variants. The three variants were identical from position 586 down to their 3′ end, but differed in their 5′ regions, suggesting alternatively spliced variants, involving the same exon-intron junction. All three differed from the sequence of ING1 described by Garkavtsev, et al., in *Nat. Genet.*, 14:415-420 (1996). These three variants are set out as SEQ ID NOS: 1, 3 and 4.

[0001] All of the sequences were then analyzed. The ORFs of SEQ ID NOS: 2, 1 and 4 (SEQ ID NO: 2 is the originally disclosed , ING1 sequence), encode polypeptides of 294, 279 and 235 amino acids, of which 233 are encoded by the 3′ region common to the three sequences. These putative sequences are set out as SEQ ID NOS: 19, 5, and 7. With

respect to SEQ ID NO: 3, however, no translational initiation site could be identified in its 5′ region.

EXAMPLE 5

[0031] The data regarding SEQ ID NO: 3, described supra, suggested further experiments to find additional ORFs in the 5-end of variant transcripts of the molecule. In order to determine this, 5'-RACE -PCR was carried out using gene specific and adapted specific primers, together with commercially available products, and standard methodologies.

[0032] The primers used for these experiments were:

CACACAGGATCCATGTTGAGTCCTGCCAACGG

CGTGGTCGTGGTTGCTGGACGCG

(SEQ ID NOS: 9 and 10), for SEQ ID NO: 1;

CCCAGCGGCCCTGACGCTGTC

CGTGGTCGTGGTTGCTGGACGCG

(SEQ ID NOS: 11 and 12), for SEQ ID NO: 3; and

GGAAGAGATAAGGCCTAGGGAAG

CGTGGTCGTGGTTGCTGGACGCG

(SEQ ID NOS: 13 and 14), for SEQ ID NO: 4.

[0033] Cloning and sequencing of the products of RACE PCR showed that the variant sequence of SEQ ID NO: 4 was 5′ to SEQ ID NO: 3, and that full length cDNA for the variant SEQ ID NO: 3 contained an additional exon 609 nucleotides long, positioned between SEQ ID NO: 3 and the shared, 3′ sequence referred to supra. This exon did not include an ORF. The first available initiation site would be an initial methionine at amino acid 70 of SEQ ID NO: 1. Thus, if expressed, SEQ ID NO: 3 would correspond to a molecule with a 681 base pair, untranslated 5′ end and a region encoding 210 amino acids (SEQ ID NO: 6).

EXAMPLE 6

[0034] The presence of transcript variants with at least 3 different transcriptional initiation sites, and possibly different promoters, suggested that mRNA expression might be under different, tissue specific regulation.

[0035] To determine this, variant-specific primers were synthesized, and RT-PCR was carried out on a panel of tissues, using standard methods.

[0036] SEQ ID NO: 1 was found to be expressed universally in all of the normal breast, brain and testis tissues examined, in six breast cancer lines, and 8 melanoma cell lines, and in cultured melanocytes. SEQ ID NO: 3 was found to be expressed in four of the six breast cancer lines, normal testis, liver, kidney, colon and brain. SEQ ID NO: 4 was only found to be expressed by normal testis cells and weakly in brain cells.

EXAMPLE 7

[0037] A further set of experiments were carried out to determine if antibodies against ING1 were present in sera of normal and cancer patients. A phase plaque immunoassay of the type described supra was carried out, using clones of SEQ ID NO: 1 as target. Of 14 allogeneic sera taken from breast cancer patients, two were positive at 1:200 dilutions. All normal sera were negative.

EXAMPLE 8

[0038] The BR11 cDNA library described supra was then screened, using SEQ ID NO: 1 and standard methodologies. A 772 base pair cDNA was identified, which was different from any sequences in the data banks consulted. The sequence of this cDNA molecule is set out at SEQ ID NO: 8.

[0039] The cDNA molecule set forth as SEQ ID NO: 1 was then used in Southern blotting experiments. In brief, genomic DNA was isolated from normal human tissue, digested with BamHI or Hind III, and then separated onto 0.7% agarose gel, blotted onto nitrocellulose filters, and hybridized using 32P labelled SEQ ID NO: 1, at high stringency conditions (aqueous buffer, 65°C). The probes were permitted to hybridize overnight, and then exposed for autoradiography. Two hybridizing DNA species were identified, i.e., SEQ ID NOS: 1 and 8.

EXAMPLE 9

[0001] The cDNA molecule set forth in SEQ ID NO: 8 was then analyzed. 5'-RACE PCR was carried out using normal fetus cDNA. Full length cDNA for the molecule is 772 base pairs long, without the poly A tail. It shows strong homology to SEQ ID NO: 1, with the strongest homology in the 5' two-thirds (76% identity over nucleotide 1-480); however, the longest ORF is only 129 base pairs, and would encode a polypeptide 42 amino acids long which was homologous to, but much shorter than, the expected expression product of SEQ ID NO: 1.

[0001] In addition to the coding region, SEQ ID NO: 8 contains 203 base pairs of 5'-untranslated region, and 439 base pairs of 3'-untranslated region.

[0042] RT-PCR assays were carried out, as described supra. All of the normal tissues tested, including brain, colon, testis, tissue and breast, were positive for expression of this gene. Eight melanoma cell lines were tested, of which seven showed varying levels of expression, and one showed no expression. Six breast cancer cell lines were tested, of which four showed various levels of expression, and two showed no expression.

EXAMPLE 10

[0043] An additional breast cancer cDNA library, referred to as "BR17-128", was screened, using autologous sera. A cDNA molecule was identified.

[0001] Analysis of the sequence suggested that it was incomplete at the 5' end. To extend the sequence, a testicular cDNA library was screened with a nucleotide probe based upon the partial sequence identified in the breast cancer library. An additional 1200 base pairs were identified following these screenings. The 2030 base pairs of information are set forth in SEQ ID NO: 15.

[0045] The longest open reading frame is 1539 base pairs, corresponding to a protein of about 59.15 kilodaltons, and 512 amino acids. The deduced amino acid sequence is set forth at SEQ ID NO: 16.

[0046] RT-PCR was then carried out using the following primers:

CACACAGGATCCATGCAGGCCCCGCACAAGGAG

CACACAAAGCTTCTAGGATTTGGCACAGCCAGAG

(SEQ ID NOS: 17 and 18)

[0047] Strong signals were observed in normal testis and breast tissue, and weak expression was observed in placenta.

[0048] No expression was found in normal brain, kidney, liver, colon, adrenal, fetal brain, lung, pancreas, prostate, thymus, uterus, and ovary tissue of tumor cell lines tested, 2 of the breast cancer lines were strongly positive and two were weakly positive. Of melanoma two of 8 were strongly positive, and 3 were weakly positive. Of lung cancer cell lines, 4 of 15 were strongly positive, and 3 were weakly positive.

[0049] When cancer tissue specimens were tested, 16 of 25 breast cancer samples were strongly positive, and 3 additional samples were weakly positive. Two of 36 melanoma samples were positive (one strong, one weak). All other cancer tissue samples were negative.

[0050] When Northern blotting was carried out, a high molecular weight smear was observed in testis, but in no other tissues tested.

EXAMPLE 11

[0051] Further experiments were carried out using the tumor sample referred to in example 10, supra. This sample was derived from a subcutaneous metastasis of a 60 year old female breast cancer patient. Total RNA was extracted, as described supra. Following the extraction, a cDNA library was constructed in λ-ZAP expression vectors, also as described supra. Screening was carried out, using the protocol set forth in example 1. A total of 7 x 105 pfus were screened. Fourteen reactive clones were identified, purified, and sequenced. The sequences were then compared to published sequences in GenBank and EST databases. These analyses indicated that the clones were derived from seven distinct genes, two of which were known, and five unknown. The two known genes were "PBK-1" (three clones), and TI-227 (one clone). These are universally expressed genes, with the libraries referred to supra showing ESTs for these genes from many different tissues.

[0052] With respect to the remaining 10 clones, six were derived from the same gene, referred to hereafter as "NY-BR-1." Three cDNA sequences were found in the EST database which shared identity with the gene. Two of these (AI 951118 and AW 373574)

were identified as being derived from a breast cancer library, while the third (AW 170035), was from a pooled tissue source.

EXAMPLE 12

[0053] The distribution of the new gene NY-BR-1 referred to supra was determined via RT-PCR. In brief, NY-BR-1 gene specific oligonucleotide primers were designed to amplify cDNA segments 300-600 base pairs in length, with primer melting temperatures estimated at 65-70°C.

[0054] The RT-PCR was then carried out over 30 amplification cycles, using a thermal cycler, and an annealing temperature of 60°C. Products were analyzed via 1.5% gel electrophoresis, and ethidium bromide visualization. Fifteen normal tissues (adrenal gland, fetal brain, lung, mammary gland, pancreas, placenta, prostate, thymus, uterus, ovary, brain, kidney, liver, colon and testis) were assayed. The NY-BR-1 clone gave a strong signal in mammary gland and testis tissue, and a very faint signal in placenta. All other tissues were negative. The other clones were expressed universally, based upon comparison to information in the EST database library, and were not pursued further.

[0055] The expression pattern of NY-BR-1 in cancer samples was then tested, by carrying out RT-PCR, as described supra, on tumor samples.

[0056] In order to determine the expression pattern, primers:

caaagcagag cctcccgaga ag

(SEQ ID NO: 20) and

cctatgctgc tcttcgattc ttcc

(SEQ ID NO: 21) were used.

[0057] Of twenty-five breast cancer samples tested, twenty two were positive for NY-BR-1. Of these, seventeen gave strong signals, and five gave weak to modest signals.

[0058] An additional 82 non-mammary tumor samples were also analyzed, divided into 36 melanoma, 26 non small cell lung cancer, 6 colon cancer, 6 squamous cell carcinoma, 6 transitional cell carcinoma, and two leiyomyosarcomas. Only two melanoma samples were positive for NY-BR-1 expression.

[0059] The study was then extended to expression of NY-BR-1 in tissue culture. Cell lines derived from breast tumor, melanoma, and small cell lung cancer were studied. Four of six breast cancer cells were positive (two were very weak), four of eight melanoma (two very weak), and seven of fourteen small cell lung cancer lines (two very weak) were positive.

EXAMPLE 13

[0060] Studies were continued in order to determine the complete cDNA sequence for NY-BR-1. First, the sequences of the six clones referred to supra were compiled using standard methods, to produce a nucleotide sequence 1464 base pairs long. Analysis of the open reading frame showed a continuous ORF throughout, indicating that the compiled sequence is not complete.

[0061] Comparison of the compiled sequence with the three EST library sequences referred to supra allowed for further extension of the sequence. The EST entry AW170035 (446 base pairs long) overlapped the compiled sequence by 89 base pairs at its 5' end, permitting extension of the sequence by another 357 base pairs. A translational terminal codon was identified in this way, leading to a molecule with a 3'-untranslated region 333 base pairs long. The 5' end of the molecule was lacking, however, which led to the experiments described infra.

EXAMPLE 14

[0062] In order to determine the missing, 5' end of the clone described supra, a commercially available testis cDNA expression library was screened, using a PCR expression product of the type described supra, as a probe. In brief, 5x104 pfus per 150 mm plate were transferred to nitrocellulose membranes, which were then submerged in denaturation solution (1.5M NaCl and 0.5 M NaOH), transferred to neutralization solution (1.5 M NaCl and 0.5M Tris-HCl), and then rinsed with 0.2M Tris-HCl, and 2xSSC. Probes were labelled with 32P and hybridization was carried out at high stringency conditions (i.e., 68°C, aqueous buffer). Any positive clones were subcloned, purified, and in vivo excised to plasmid PBK-CMV, as described supra.

[0063] One of the clones identified in this way included an additional 1346 base pairs at the 5' end; however, it was not a full length molecule. A 5'-RACE-PCR was carried out, using commercially available products. The PCR product was cloned into plasmid vector pGEMT and sequenced. The results indicated that cDNA sequence extended 1292 base pairs further, but no translation initiation site could be determined, because no stop codons could be detected. It could be concluded, however, that the cDNA of the NY-BR17 clone comprises at least 4115 nucleotides, which are presented as SEQ ID NO: 22. The molecule, as depicted, encodes a protein at least about 152.8 kDA in molecular weight. Structurally, there are 99 base pairs 5' to the presumed translation initiation site, and an untranslated segment 333 base pairs long at the 3' end. The predicted amino acid sequence of the coding region for SEQ ID NO: 22 is set out at SEQ. ID NO: 23.

[0064] SEQ ID NO: 23 was analyzed for motifs, using the known search programs PROSITE and Pfam. A bipartite nuclear localization signal motif was identified at amino acids 17-34, suggesting that the protein is a nuclear protein. Five tandem ankyrin repeats were identified, at amino acids 49-81, 82-114, 115-147, 148-180 and 181-213. A bZIP site (i.e., a DNA binding site followed by a leucine zipper motif) was found at amino acid positions 1077-1104, suggesting a transcription factor function. It was also observed that three repetitive elements were identified in between the ankyrin repeats and the bZIP DNA binding site. To elaborate, a repetitive element 117 nucleotides long is trandemly repeated 3 times, between amino acids 459-815. The second repetitive sequence, consisting of 11 amino acids, repeats 7 times between amino acids 224 and 300. The third repetitive element, 34 amino acids long, is repeated twice, between amino acids 301-368.

EXAMPLE 15

[0065] The six clones described supra were compared, and analysis revealed that they were derived from two different splice variants. Specifically, two clones, referred to as "BR17-8" and "BR 17-44a", contain one more exon, of 111 base pairs (nucleotides 3015-3125 of SEQ ID NO: 22), which encodes amino acids 973-1009 of SEQ ID NO: 23, than do clones BR 17-1a, BR17-35b and BR17-44b. The shortest of the six clones, BR17-128, starts 3' to the additional exons. The key structural elements referred to supra were present in both splice variants, suggesting that there was no difference in biological function.

[0066] The expression pattern of the two splice variants was assessed via PT-PCR, using primers which spanned the 111 base pair exon referred to supra.

[0067] The primers used were:

aatgggaaca agagctctgc ag

(SEQ ID NO: 24) and

gggtcatctg aagttcagca ttc

(SEQ ID NO: 25)

[0068] Both variants were expressed strongly in normal testis and breast. The longer variant was dominant in testis, and the shorter variant in breast cells. When breast cancer cells were tested, co-typing of the variant was observed, (7 strongly, 2 weakly positive, and 1 negative), with the shorter variant being the predominant form consistently.

EXAMPLE 16

[0069] The frequency of antibody response against NY-BR-1 in breast cancer patients was tested. To do this, a recombinant protein consisting of amino acids 993-1188 of SEQ ID NO: 23 was prepared. (This is the protein encoded by clone BR 17-128, referred to supra). A total of 140 serum samples were taken from breast cancer patients, as were 60 normal serum samples. These were analyzed via Western blotting, using standard methods.

[0070] Four of the cancer sera samples were positive, including a sample from patient BR17. All normal sera were negative.

[0071] An additional set of experiments was then carried out to determine if sera recognized the portion of NY-BR-1 protein with repetitive elements. To do this, a different recombinant protein, consisting of amino acids 405-1000 was made, and tested in Western blot assays. None of the four antibody positive sera reacted with this protein indicating that an antibody epitope is located in the non-repetitive, carboxy terminal end of the molecule.

EXAMPLE 17

[0072] The screening of the testicular cDNA library referred to supra resulted, inter alia, in the identification of a cDNA molecule that was homologous to NY-BR-1. The molecule is 3673 base pairs in length, excluding the poly A tail. This corresponded to

nucleotides 1-3481 of SEQ ID NO: 22, and showed 62% homology thereto. No sequence identity to sequences in libraries was noted. ORF analysis identified an ORF from nucleotide 641 through the end of the sequence, with 54% homology to the protein sequence of SEQ. ID NO: 23. The ATG initiation codon of this sequence is 292 base pairs further 3' to the presumed initiation codon of NY-BR-1, and is preceded by 640 untranslated base pairs at its 5' end. This 640 base pair sequence includes scattered stop codons. The nucleotide sequence and deduced amino acid sequence are presented as SEQ ID NOS: 26 and 27, respectively.

[0073] RT-PCR analysis was carried out in the same way as is described supra, using primers:

tctcatagat gctggtgctg atc

(SEQ ID NO: 28) and

cccagacatt gaattttggc agac

(SEQ ID NO: 29).

[0074] Tissue restricted mRNA expression was found. The expression pattern differed from that of SEQ ID NO: 22. In brief, of six normal tissues examined, strong signals were found in brain and testis only. There was no or weak expression in normal breast tissues, and kidney, liver and colon tissues were negative. Eight of ten 10 breast cancer specimens tested supra were positive for SEQ. ID NO: 26. Six samples were positive for both SEQ. ID NO: 22 and 26, one for SEQ. ID NO: 22 only, two for the SEQ. ID NO: 26 only, and one was negative for both.

EXAMPLE 18

[0075] Recently, a working draft of the human genome sequence was released. This database was searched, using standard methods, and NY-BR-1 was found to have sequence identity with at least three chromosome 10 clones, identified by Genbank accession numbers AL157387, AL37148, and AC067744. These localize NY-BR-1 to chromosome 10 p11.21-12.1.

[0076] The comparison of NY-BR-1 and the human genomic sequence led to definition of the exon-intron organization of NY-BR-1. In brief, the coding region of the

gene contains essentially 19 structurally distinct exons with at least 2 exons encoding 3' untranslated regions. Detailed exon-intron junction information is described at Genbank AF 269081.

[0077] The six ankyrin repeats, referred to supra, are all found within exon 7. The 357 nucleotide repeating unit is composed of exons 10-15. The available genomic sequences are not complete, however, and only one of the three copies was identified, suggesting that DNA sequences between exons 5 and 10 may be duplicated and inserted in tandem, during genetic evolution. In brief, when the isolated NY-BR-1 cDNA clone was analyzed, three complete and one incomplete copy of the repeating units were found. The exon sequences can be expressed as exons 1-2-3-4-5-6-7-8-9-(10-11-12-13-14-15)-(10A-11A-12A-13A-14A-15A)-(10B-11B-12B-13B-14B-15B)-(10C-11C-12C-13C-14C)-16-17-18-19-20-21, wherein A, B & C are inexact copies of exon 10-15 sequences. Cloned, NY-BR-1 cDNA has 38 exons in toto.

[0078] It was noted, supra, that the sequence of NY-BR-1 cDNA was not complete at the 5' end. A genomic sequence (Genbank AC067744), permitted extension of the 5' end. This extended sequence is set forth in SEQ ID NO: 31. It consists of 4194 base pairs of coding sequence, plus a 2088 base pair segment 3' to the coding segment, which is untranslated. (This excludes the poly A tail). As remarked upon previously, this sequence contains a bipartite nuclear localization signal, 5 ankyrin repeats, and a b zip site. Translation of the 5' genomic sequence led to the identification of a new translation initiation site, 168 base pairs upstream of the previously predicted ATG initiation codon. This resulted in an NY-BR-1 polypeptide including 1397 amino acids which is 56 amino acid residues longer, at the N-terminus, as compared to SEQ ID NO: 23. The additional amino acids are: MEEISAAAVKVVPGPERPSPFSQLVYTSNDSYIVHSGDLRKIHKAASRGQVRKLEK (SEQ ID NO: 30). These amino acids are positioned N-terminal to SEQ ID NO: 23, in SEQ ID NO: 32.

EXAMPLE 19

[0079] Reference was made, supra, to the two difference splice variants of NY-BR-1. Comparison of the splice variants with the genomic sequence confirmed that an alternate

splicing event, with the longer variant incorporating part of intron 33 into exon 34 (i.e., exon 17 of the basic exon/intron framework described supra), had occurred.

[0080] Key structural elements that were predicted in NY-BR-1, described supra, are present in both variants, suggesting that there is no difference in biological function, or subcellular location.

EXAMPLE 20

[0081] As with NY BR-1, the variant NY-BR-1.1, described supra, was screened against the working draft of the human genomic sequence. One clone was found with sequence identity, i.e., GenBank AL359312, derive from chromosome 9. Thus, NY-BR-1 and NY-BR-1.1 both appear to be functioning genes, on two different chromosomes. The Genbank sequences referred to herein does not contain all of NY-BR-1.1, which precludes defining exon-intron structure. Nonetheless, at least 3 exons can be defined, which correspond to exons 16-18 of the NY-BR-1 basic framework. Exon-intron junctions are conserved.

EXAMPLE 21

[0082] A series of peptides were synthesized, based upon the amino acid sequence of NY-BR-1, as set forth in SEQ ID NO: 23 and the concatenation of SEQ ID NOS: 30 & 23, as described supra and set forth at SEQ ID NO: 32. These were then tested for their ability to bind to HLA-A2 molecules and to stimulate CTL proliferation, using an ELISPOT assay. This assay involved coating 96-well, flat bottom nitrocellulose plates with 5ug/ml of anti-interferon gamma antibodies in 100 ul of PBS per well, followed by overnight incubation. Purified CD8+ cells, which had been separated from PBL samples via magnetic beads coated with anti-CD8 antibodies were then added, at 1x105 cells/well, in RPMI 1640 medium, that had been supplemented with 10% human serum, L-asparagine (50 mg/l), L-arginine (242 mg/l), L-glutamine (300 mg/l), together with IL-2 (2.5ng/ml), in a final volume of 100 ul. CD8+ effector cells were prepared by presensitizing with peptide, and were then added at from 5x103 to 2x104 cells/well. Peptides were pulsed onto irradiated T2 cells at a concentration of 10ug/ml for 1 hour, washed and added to effector cells, at 5x104 cells/well. The plates were incubated for 16 hours at 37°C, washed six times with 0.05% Tween

20/PBS, and were then supplemented with biotinylated, anti-interferon gamma specific antibody at 0.5 ug/ml. After incubation for 2 hours at 37°C, plates were washed, and developed with commercially available reagents, for 1 hour, followed by 10 minutes of incubation with dye substrate. Plates were then prepped for counting, positives being indicated by blue spots. The number of blue spots/well was determined as the frequency of NY-ESO-1 specific CTLs/well.

[0083] Experiments were run, in triplicate, and total number of CTLs was calculated. As controls, one of reagents alone, effector cells alone, or antigen presenting cells alone were used. The difference between the number of positives in stimulated versus non-stimulated cells, was calculated as the effective number of peptide specific CTLs above background. Three peptides were found to be reactive, i.e.:

LLSHGAVIEV (amino acids 102-111 of SEQ ID NO: 23, 158-167 of SEQ ID NO: 32)

SLSKILDTV (amino acids 904-912 of SEQ ID NO: 23, 960-968 of SEQ ID NO: 32)

SLDQKLFQL (amino acids 1262-1270 of SEQ ID NO: 23, 1318-1326 of SEQ ID NO: 32).

[0084] The complete list of peptides tested, with reference to their position in SEQ ID NO: 23, follows:

Peptide	Position
FLVDRKVCQL	35-43
ILIDSGADI	68-76
AVYSEILSV	90-98
ILSVVAKLL	95-103
LLSHGAVIEV	102-111
KLLSHGAVI	101-109
FLLIKNANA	134-142
MLLQQNVDV	167-175
GMLLQQNVDV	166-175
LLQQNVDVFA	168-177

IAWEKKETPV	361-370
SLFESSAKI	430-438
CIPENSIYQKV	441-450
KVMEINREV	449-457
ELMDMQTFKA	687-696
ELMDMQTFKA	806-815
SLSKILDTV	904-912
KILDTVHSC	907-915
ILNEKIREEL	987-996
RIQDIELKSV	1018-1027
YLLHENCML	1043-1051
CMLKKEIAML	1049-1058
AMLKLELATL	1056-1065
KILKEKNAEL	1081-1090
VLIAENTML	1114-1122
CLQRKMNVDV	1174-1183
KMNVDVSST	1178-1186
SLDQKLFQL	1262-1270
KLFQLQSKNM	1266-1275
FQLQSKNMWL	1268-1277
QLQSKNMWL	1269-1277
NMWLQQQLV	1274-1282
WLQQQLVHA	1276-1284
KITIDIHFL	1293-1301

EXAMPLE 22

[0085] Expression of the full length NY-BR-1 molecule was analyzed, by determining the presence of mRNA, in various normal and tumor tissue samples.

[0086] RT-PCR assays were carried out, as described in examples 5 & 9, on a variety of tissue samples.

[0087] Expression on the mRNA level was found in normal breast and testis tissue, but in none of normal adrenal gland, fetal brain, lung, pancreatic, placental, prostate, thymus, uterine, ovarian, adult brain, kidney, liver or colon tissue.

[0088] With respect to cancer tissue samples, 19/34 breast cancer samples were positive, as were 9/34 prostate cancer biopsies.

EXAMPLE 23

[0089] These experiments describe work which identified and verified two, naturally processed T cell epitopes that consist of amino acid sequences found in NY-BR-1.

[0090] Sequences encoding NY-BR-1 were excised from plasmid pQE9, via standard restriction enzyme digestion, and were cloned into BamHI-Hind III sites of commercially available plasmid pcDNA31(-).

[0091] The resulting vectors were then transfected into COS-7 cells. To accomplish this, $2x10^4$COS-7 cells were admixed with 150 ng of the construct described supra, and 150 ng of plasmid pcDNA-AmpI, which contained cDNA encoding HLA-A2. The standard DE AE-dextran chloroquine method was used. Transfectants were then incubated at 37°C for 48 hours, and then tested in a T cell stimulation assay, after 24 hours, as described infra.

[0092] The transfectants were tested to determine if they could stimulate production of TNF-α by CTLs specific for complexes of HLA-A2 molecules and one of the peptides described supra. The CTLs used were CD8+ T cell clones. "NW1100-CTL-7," "NW1100-CTL39," and "NW1100-CTL43." These three CD8+ T cell clones had been generated via repeated in vitro stimulation with either LLSHGAVIEV or SLSKILDTV, using standard methods.

[0093] To test if the transfectants stimulated the CD8+ cells, 5000 of these CD8+ cells, in 100 µl RPMI supplemented with 10% human serum, and 25 U/ml of recombinant human IL-2 were added to micowells containing the transfectants. After 24 hours, 50 µl samples of supernatant were collected, and TNFα content was determined by testing cytotoxicity against WEHI 164 clone 13 cells, in an MTT colorimetric assay, which is a standard method for showing TNFα production.

[0094] The results are shown in figures 1, 2 and 3. Briefly, both peptide/HLA-A2 complexes were recognized by CD8+ T cells obtained from breast cancer patient identified as NW-1100. These results indicate that the two peptides are, in fact, naturally processed.

[0095] The foregoing examples describe the isolation of a nucleic acid molecule which encodes a cancer associated antigen. "Associated" is used herein because while it is clear that the relevant molecule was expressed by several types of cancer, other cancers, not screened herein, may also express the antigen.

[0096] The invention relates to nucleic acid molecules which encode the antigens encoded by, e.g., SEQ ID NOS: 1, 3, 8, 15, 22, 26 and 31 as well as the antigens encoded thereby, such as the proteins with the amino acid sequences of SEQ ID NOS: 5, 6, 7, 16, 23, 27, 30 and 32. It is to be understood that all sequences which encode the recited antigen are a part of the invention. Also a part of the invention are those nucleic acid molecules which have complementary nucleotide sequences which hybridize to the referred sequences, under stringent conditions. "Stringent conditions" as used herein refers, e.g., to prehybridization in 6xSSC/0.05 BLOTTO for 2 hours, followed by adding a probe mixed with salmon sperm DNA and overnight incubation at 68°C, followed by two one minute washes with 2xSSC/0.2% room temperature, and then three twenty minute washes with 2xSSC/0.2% SDS (68°C). An optional additional one or two high stringency washes with 0.2xSSC/0.2% SDS, for 20 minutes, at 68°C, may be included.

[0097] Also a part of the invention are proteins, polypeptides, and peptides, which comprise, e.g., at least nine consecutive amino acids found in SEQ ID NO: 23 or 32, or at least nine consecutive amino acids of the amino acids of SEQ ID NO: 30 or 32. Proteins, polypeptides and peptides comprising nine or more amino acids of SEQ ID NO: 5, 6, 7, 16 or 27 are also a part of the invention. Especially preferred are peptides comprising or consisting of amino acids 102-111, 904-912, or 1262-1270 of SEQ ID NO: 23, which are paralleled in SEQ ID NO: 32. Such peptides may, but do not necessarily provoke CTL responses when complexed with an HLA molecule, such as an HLA-A2 molecule. They may also bind to different MHC or HLA molecules, including, but not being limited to, HLA-A1, A2, A3, B7, B8, Cw3, Cw6, or serve, e.g., as immunogens, as part of immunogenic cocktail compositions, where they are combined with other proteins or polypeptides, and so forth. Also a part of the invention are the nucleic acid molecules which encode these molecules,

such as "minigenes," expression vectors that include the coding regions, recombinant cells containing these, and so forth. All are a part of the invention.

[0098] Also a part of the invention are expression vectors which incorporate the nucleic acid molecules of the invention, in operable linkage (i.e., "operably linked") to a promoter. Construction of such vectors, such as viral (e.g., adenovirus or Vaccinia virus) or attenuated viral vectors is well within the skill of the art, as is the transformation or transfection of cells, to produce eukaryotic cell lines, or prokaryotic cell strains which encode the molecule of interest. Exemplary of the host cells which can be employed in this fashion are COS cells, CHO cells, yeast cells, insect cells (e.g., Spodoptera frugiperda), NIH 3T3 cells, and so forth. Prokaryotic cells, such as E. coli and other bacteria may also be used. Any of these cells can also be transformed or transfected with further nucleic acid molecules, such as those encoding cytokines, e.g., interleukins such as IL-2, 4, 6, or 12 or HLA or MHC molecules.

[0099] Also a part of the invention are the antigens described herein, both in original form and in any different post translational modified forms. The molecules are large enough to be antigenic without any posttranslational modification, and hence are useful as immunogens, when combined with an adjuvant (or without it), in both precursor and post-translationally modified forms. Antibodies produced using these antigens, both poly and monoclonal, are also a part of the invention as well as hybridomas which make monoclonal antibodies to the antigens. The whole protein can be used therapeutically, or in portions, as discussed infra. Also a part of the invention are antibodies against this antigen, be these polyclonal, monoclonal, reactive fragments, such as Fab, $(F(ab)_2$, and other fragments, as well as chimeras, humanized antibodies, recombinantly produced antibodies, and so forth.

[00100] As is clear from the disclosure, one may use the proteins and nucleic acid molecules of the invention diagnostically. The SEREX methodology discussed herein is premised on an immune response to a pathology associated antigen. Hence, one may assay for the relevant pathology via, e.g., testing a body fluid sample of a subject, such as serum, for reactivity with the antigen per se. Reactivity would be deemed indicative of possible presence of the pathology. So, too, could one assay for the expression of any of the antigens via any of the standard nucleic acid hybridization assays which are well known to the art, and

need not be elaborated upon herein. One could assay for antibodies against the subject molecules, using standard immunoassays as well.

[00101] As was shown in, e.g., examples 22 & 23, the invention relates in particular to methods for determining if cancer is present, such as breast cancer or pancreatic cancer, by assaying for expression of NY-BR-1, as defined supra, via a nucleotide based assay, such as polymerase chain reaction (PCR) or some other form of nucleotide hybridization assay, a protein based assay, such as an immunoassay, or a peptide based assay where one either looks for, or utilizes, CD8+ cells which react specifically with complexes of peptides and their partner HLA molecule, such as LLSHGAVIEV or SLSKILDTV, and HLA-A2. As with the nucleotide and protein based assays, these peptide based assays are especially useful in determining breast and/or pancreatic cancer. The assays of the invention, in all forms, can be used to determine presence, progression, and/or regression of cancer, such as breast and/or pancreatic cancer, and can then be used to determine the efficacy of therapeutic regimes, especially when the regime is directed against breast and/or pancreatic cancer.

[00102] Analysis of SEQ ID NO: 1, 3, 4, 8, 15, 22, 26 and 31 will show that there are 5′ and 3′ non-coding regions presented therein. The invention relates to those isolated nucleic acid molecules which contain at least the coding segment, and which may contain any or all of the non-coding 5′ and 3′ portions.

[00103] Also a part of the invention are portions of the relevant nucleic acid molecules which can be used, for example, as oligonucleotide primers and/or probes, such as one or more of SEQ ID NOS: 9, 10, 11, 12, 13, 14, 17, 18,20, 21, 24, 25, 28, and 29 as well as amplification products like nucleic acid molecules comprising at least nucleotides 305-748 of SEQ ID NO: 1, or amplification products described in the examples, including those in examples 12, 14, etc.

[00104] As was discussed supra, study of other members of the "CT" family reveals that these are also processed to peptides which provoke lysis by cytolytic T cells. There has been a great deal of work on motifs for various MHC or HLA molecules, which is applicable here. Hence, a further aspect of the invention is a therapeutic method, wherein one or more peptides derived from the antigens of the invention which bind to an HLA molecule on the surface of a patient's tumor cells are administered to the patient, in an amount sufficient for the peptides to bind to the MHC/HLA molecules, and provoke lysis by T cells. Any

combination of peptides may be used. These peptides, which may be used alone or in combination, as well as the entire protein or immunoreactive portions thereof, may be administered to a subject in need thereof, using any of the standard types of administration, such as intravenous, intradermal, subcutaneous, oral, rectal, and transdermal administration. Standard pharmaceutical carriers, adjuvants, such as saponins, GM-CSF, and interleukins and so forth may also be used. Further, these peptides and proteins may be formulated into vaccines with the listed material, as may dendritic cells, or other cells which present relevant MHC/peptide complexes.

[00105] Of particular interest, are peptides shown to be natural epitopes of the NY-BR-1 molecule, such as LLSHGAVIEV and SLSKILDTV. By "natural epitopes" is meant that CD8+ cells taken from patients with cancer recognize and lyse cells which present these peptides on their surface. It is more desirable to use peptides which have been shown to be naturally occurring epitopes in an in vivo context, because these peptides can lead to expansion of pre-existing populations of relevant CD8+ cells. In parallel, CD8+ cells which are specific to the complexes can be used therapeutically. Hence, in any of the therapeutic approaches discussed herein relating to peptides or minigenes , it is especially preferred to use one or both of these peptide sequences, or minigenes which encode them.

[00106] Similarly, the invention contemplates therapies wherein nucleic acid molecules which encode the proteins of the invention, one or more or peptides which are derived from these proteins are incorporated into a vector, such as a Vaccinia or adenovirus based vector, to render it transfectable into eukaryotic cells, such as human cells. Nucleic acid molecules which encode one or more of the peptides may be incorporated into these vectors, which are then the major constituent of nucleic acid bases therapies.

[00107] Any of these assays can also be used in progression/regression studies. One can monitor the course of abnormality involving expression of these antigens simply by monitoring levels of the protein, its expression, antibodies against it and so forth using any or all of the methods set forth supra.

[00108] It should be clear that these methodologies may also be used to track the efficacy of a therapeutic regime. Essentially, one can take a baseline value for a protein of interest using any of the assays discussed supra, administer a given therapeutic agent, and

then monitor levels of the protein thereafter, observing changes in antigen levels as indicia of the efficacy of the regime.

[00109] As was indicated supra, the invention involves, inter alia, the recognition of an "integrated" immune response to the molecules of the invention. One ramification of this is the ability to monitor the course of cancer therapy. In this method, which is a part of the invention, a subject in need of the therapy receives a vaccination of a type described herein. Such a vaccination results, e.g., in a T cell response against cells presenting HLA/peptide complexes on their cells. The response also includes an antibody response, possibly a result of the release of antibody provoking proteins via the lysis of cells by the T cells. Hence, one can monitor the effect of a vaccine, by monitoring an antibody response. As is indicated, supra, an increase in antibody titer may be taken as an indicia of progress with a vaccine, and vice versa. Hence, a further aspect of the invention is a method for monitoring efficacy of a vaccine, following administration thereof, by determining levels of antibodies in the subject which are specific for the vaccine itself, or a large molecule of which the vaccine is a part.

[00110] The identification of the subject proteins as being implicated in pathological conditions such as cancer also suggests a number of therapeutic approaches in addition to those discussed supra. The experiments set forth supra establish that antibodies are produced in response to expression of the protein. Hence, a further embodiment of the invention is the treatment of conditions which are characterized by aberrant or abnormal levels of one or more of the proteins, via administration of antibodies, such as humanized antibodies, antibody fragments, and so forth. These may be tagged or labelled with appropriate cystostatic or cytotoxic reagents.

[00111] T cells may also be administered. It is to be noted that the T cells may be elicited in vitro using immune responsive cells such as dendritic cells, lymphocytes, or any other immune responsive cells, and then reperfused into the subject being treated.

[00112] Note that the generation of T cells and/or antibodies can also be accomplished by administering cells, preferably treated to be rendered non-proliferative, which present relevant T cell or B cell epitopes for response, such as the epitopes discussed supra.

[00113] The therapeutic approaches may also include antisense therapies, wherein an antisense molecule, preferably from 10 to 100 nucleotides in length, is administered to the

subject either "neat" or in a carrier, such as a liposome, to facilitate incorporation into a cell, followed by inhibition of expression of the protein. Such antisense sequences may also be incorporated into appropriate vaccines, such as in viral vectors (e.g., Vaccinia), bacterial constructs, such as variants of the known BCG vaccine, and so forth.

[00114] Also a part of this invention are antibodies, e.g., polyclonal and monoclonal, and antibody fragments e.g., single chain Fv, Fab, diabodies etc., that specifically bind the peptides or HLA/peptide complexes disclosed herein. Preferably the antibodies, the antibody fragments and T cell receptors bind the HLA/peptide complexes in a peptide-specific manner. Such antibodies are useful, for example, in identifying cells presenting the HLA/peptide complexes, particularly complexes comprising an HLA-A1, A2, A3, A26, HLA-B7, B8, B15, B27, B35, B44, B51, B57, Cw3, or Cw6 molecule, preferably HLA-A2 or B57, and a peptide consisting essentially of the sequences described supra, such as amino acids 102-111, 904-912, or 1262-1270 of SEQ ID NO: 23.

[00115] Such antibodies are also useful in promoting the regression or inhibiting the progression of a tumor which expresses complexes of the HLA and peptide. Polyclonal antisera and monoclonal antibodies specific to the peptides or HLA/peptide complexes of this invention may be generated according to standard procedures. See e.g., Catty, D., *Antibodies, A Practical Approach*, Vol. 1, IRL Press, Washington DC (1988); Klein, *J. Immunology: The Science of Cell-Non-Cell Discrimination,* John Wiley and Sons, New York (1982); Kennett, R., et al., *Monoclonal Antibodies, Hybridoma, A New Dimension In Biological Analyses*, Plenum Press, New York (1980); Campbell, A., *Monoclonal Antibody Technology, in Laboratory Techniques and Biochemistry and Molecular Biology*, Vol. 13 (Burdon, R., et al. EDS.), Elsevier Amsterdam (1984); Eisen, H. N., *Microbiology*, third edition, Davis, B. D., et al. EDS. (Harper & Rowe, Philadelphia (1980); Kohler and Milstein, *Nature*, 256:495 (1975) all incorporated herein by reference.) Methods for identifying Fab molecules endowed with the antigen-specific, HLA-restricted specificity of T cells has been described by Denkberg, et al., *Proc. Natl. Acad. Sci.*, 99:9421-9426 (2002) and Cohen, et al., *Cancer Research*, 62:5835-5844 (2002) (both incorporated herein by reference). Methods for generating and identifying other antibody molecules, e.g., scFv and diabodies are well known in the art, see e.g., Bird, et al., *Science*, 242:423-426 (1988); Huston, et al., *Proc. Natl. Acad. Sci.*, 85:5879-5883 (1988); Mallender and Voss, *J. Biol. Chem.*, 269:199-206

(1994); Ito and Kurosawa, *J. Biol. Chem.*, 27:20668-20675 (1993), and; Gandecha, et al., *Prot. Express Purif.*, 5:385-390 (1994)(all incorporated herein by reference).

[00116] The antibodies of this invention can be used for experimental purposes (e.g. localization of the HLA/peptide complexes, immunoprecipitations, Western blots, flow cytometry, ELISA etc.) as well as diagnostic or therapeutic purposes, e.g., assaying extracts of tissue biopsies for the presence of HLA/peptide complexes, targeting delivery of cytotoxic or cytostatic substances to cells expressing the appropriate HLA/peptide complex. The antibodies of this invention are useful for the study and analysis of antigen presentation on tumor cells and can be used to assay for changes in the HLA/peptide complex expression before, during or after a treatment protocol, e.g., vaccination with peptides, antigen presenting cells, HLA/peptide tetramers, adoptive transfer or chemotherapy. The antibodies and antibody fragments of this invention may be coupled to diagnostic labeling agents for imaging of cells and tissues that express the HLA/peptide complexes or may be coupled to therapeutically useful agents by using standard methods well-known in the art. The antibodies also may be coupled to labeling agents for imaging e.g., radiolabels or fluorescent labels, or may be coupled to, e.g., biotin or antitumor agents, e.g., radioiodinated compounds, toxins such as ricin, methotrexate, cytostatic or cytolytic drugs, etc. Examples of diagnostic agents suitable for conjugating to the antibodies of this invention include e.g., barium sulfate, diatrizoate sodium, diatrizoate meglumine, iocetamic acid, iopanoic acid, ipodate calcium, metrizamide, tyropanoate sodium and radiodiagnostics including positron emitters such as fluorine-18 and carbon-11, gamma emitters such as iodine-123, technitium-99m, iodine-131 and indium-111, nuclides for nuclear magnetic resonance such as fluorine and gadolinium. As used herein, "therapeutically useful agents" include any therapeutic molecule which are preferably targeted selectively to a cell expressing the HLA/peptide complexes, including antineoplastic agents, radioiodinated compounds, toxins, other cytostatic or cytolytic drugs. Antineoplastic therapeutics are well known and include: aminoglutethimide, azathioprine, bleomycin sulfate, busulfan, carmustine, chlorambucil, cisplatin, cyclophosphamide, cyclosporine, cytarabidine, dacarbazine, dactinomycin, daunorubicin, doxorubicin, taxol, etoposide, fluorouracil, interferon-.alpha., lomustine, mercaptopurine, methotrexate, mitotane, procarbazine HCl, thioguanine, vinblastine sulfate and vincristine sulfate. Additional antineoplastic agents include those disclosed in Chapter 52, Antineoplastic

Agents (Paul Calabresi and Bruce A. Chabner), and the introduction thereto, 1202-1263, of Goodman and Gilman's "The Pharmacological Basis of Therapeutics", Eighth Edition, 1990, McGraw-Hill, Inc. (Health Professions Division). Toxins can be proteins such as, for example, pokeweed anti-viral protein, cholera toxin, pertussis toxin, ricin, gelonin, abrin, diphtheria exotoxin, or Pseudomonas exotoxin. Toxin moieties can also be high energy-emitting radionuclides such as ^{131}I, ^{90}Y or any other alpha, beta and auger emitting that are known within the art. The antibodies may be administered to a subject having a pathological condition characterized by the presentation of the HLA/peptide complexes of this invention, e.g., melanoma or other cancers, in an amount sufficient to alleviate the symptoms associated with the pathological condition.

[00117] Soluble T cell receptors (TcR) which specifically bind to the HLA/peptide complexes described herein are also an aspect of this invention. In their soluble form T cell receptors are analogous to a monoclonal antibody in that they bind to HLA/peptide complex in a peptide-specific manner. Immobilized TcRs or antibodies may be used to identify and purify unknown peptide/HLA complexes which may be involved in cellular abnormalities. Methods for identifying and isolating soluble TcRs are known in the art, see for example WO 99/60119, WO 99/60120 (both incorporated herein by reference) which describe synthetic multivalent T cell receptor complex for binding to peptide-MHC complexes. Recombinant, refolded soluble T cell receptors are specifically described. Such receptors may be used for delivering therapeutic agents or detecting specific peptide-MHC complexes expressed by tumor cells. WO 02/088740 (incorporated by reference) describes a method for identifying a substance that binds to a peptide-MHC complex. A peptide-MHC complex is formed between a predetermined MHC and peptide known to bind to such predetermined MHC. The complex is then use to screen or select an entity that binds to the peptide-MHC complex such as a T cell receptor. The method could also be applied to the selection of monoclonal antibodies that bind to the predetermined peptide-MHC complex.

[00118] Also a part of this invention are nucleic acid molecules encoding the antibodies and T cell receptors of this invention and host cells, e.g., human T cells, transformed with a nucleic acid molecule encoding a recombinant antibody or antibody fragment, e.g., scFv or Fab, or a TcR specific for a predesignated HLA/peptide complex as described herein, particularly a complex wherein the HLA molecule is an HLA-A1, A2, A3,

A26, HLA-B7, B8, B15, B27, B35, B44, B51, B57, Cw3 or Cw6 molecule, preferably HLA-A2 or B57, and the peptide is encoded by nucleotide sequence corresponding to a nucleotide sequence found in SEQ ID NO: 31.

[00119] Recombinant Fab or TcR specific for a predesignated HLA/peptide complex in T cells have been described in, e.g., Willemsen, et al., "A phage display selected Fab fragment with MHC class I-restricted specificity for MAGE-A1 allows for retargeting of primary human T lymphocytes" *Gene Ther.*, 2001 Nov;8(21):1601-8. and Willemsen, et al., "Grafting primary human T lymphocytes with cancer-specific chimeric single chain and two chain TCR". *Gene Ther.*, 2000 Aug; 7(16):1369-77. (both incorporated herein by reference) and have applications in an autologous T cell transfer setting. The autologous T cells, transduced to express recombinant antibody or TcR, may be infused into a patient having an pathological condition associated with cells expressing the HLA/peptide complex. The transduced T cells are administered in an amount sufficient to inhibit the progression or alleviate at least some of the symptoms associated with the pathological condition.

[00120] This invention also relates to a method for promoting regression or inhibiting progression of a tumor in a subject in need thereof wherein the tumor expresses a complex of HLA and peptide. The method comprises administering an antibody, antibody fragment or soluble T cell receptor, which specifically binds to the HLA/peptide complex, or by administering cells transduced so that they express those antibodies or TcR in amounts that are sufficient to promote the regression or inhibit progression of the tumor expressing the HLA/peptide complex, e.g., a melanoma or other cancer. Preferably the HLA is an HLA-A2, or B57 and the peptide is an NY-BR-1 derived peptide preferably a peptide consisting of the sequences set forth supra, such as amino acids 102-111, 904-912, or 1262-1270 of SEQ ID NO: 23.

[00121] The antibodies, antibody fragments and soluble T cell receptors may be conjugated with, or administered in conjunction with, an antineoplastic agent, e.g., radioiodinated compounds, toxins such as ricin, methotrexate, or a cytostatic or cytolytic agent as discussed supra. See e.g., Pastan, et al., *Biochem. Biophys. Acta.*, 133:C1-C6 (1997), Lode, et al., *Immunol. Res.*, 21:279-288 (2000) and Wihoff, et al., *Curr. Opin. Mo. Ther.*, 3:53-62 (2001) (all incorporated herein by reference) for a discussion of the

construction of recombinant immunotoxins, antibody fusions with cytokine molecules and bispecific antibody therapy or immunogene therapy.

[00122] Other features and applications of the invention will be clear to the skilled artisan, and need not be set forth herein. The terms and expression which have been employed are used as terms of description and not of limitation, and there is no intention in the use of such terms and expression of excluding any equivalents of the features shown and described or portions thereof, it being recognized that various modifications are possible within the scope of the invention.